```
Db    87 GYDPYDFFDLGEYNQKGTIETRFGGSKQELINMINTAHAYGIKV         ILEWN 146

Qy   122 PFVNDYTWTDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWASQ 181
         ||| |||||||||||||||||||||||||||:|      ||||| ||  :||||:||||:
Db   147 PFVGDYTWTDFSKVASGKYTANYLDFHPNEVKCCDBGTFGGFPDIAHEKEWDQHWLWASD 206

Qy   182 ESYAAYLRSIGIDAWRFPDYVKGYAPWVVKDWLNWWGGWAVGEYWDTNVDAVLNWAYSSGA 241
         |||||||||| ||||| |||||| |:|||||||||||||||||||||||||: ||||||||
Db   207 ESYAAYLRSIGVDAWRFDYVKGYGAWVVKDWLNWWGGWAVGEYWDTNVDALLNWAYSSGA 266

Qy   242 KVFDFALYYKMDEAFDNKNIPALVSALQNGQTVVSRDPFKAVTFVANHDTDIIWNKYLAY 301
         ||||| |||||||||| ||||||||||||||||||||||||||||||||||||||||| |
Db   267 KVFDFPLYYKMDEAFDNTNIPALVSALQNGQTVVSRDPFKAVTFVANHDTDIIWNKYLAY 326

Qy   302 AFILTYEGQPTIFYRDYEEWLNKDKLKNLIWIHENLAGGSTDIVYYDNDELIFVRNGYGD 361
         |||||||||| ||||||||||||| :||||||:|||||||:|||||:||:|||||||||
Db   327 AFILTYEGQPVIFYRDYEEWLNKDRLNNLIWIHDHLAGGSTSIVYYDSDEMIFVRNGYGS 386

Qy   362 KPGLITYINLGSSKAGRWVTVPKFPAGACIHEYTGNLGGWVDKYVYSSGWVLEAPAYDPA 421
         ||||||||||||||:||||:|||||||||||||||||||||||||| ||||||||||||||
Db   387 KPGLITYINLGSSKVGRWVYVPKFPAGACIHEYTGNLGGWVDKYVYSSGWVYLEAPAYDPA 446

Qy   422 NGQYGYSVWSYCGVG 436
         |||||||||||||||
Db   447 NGQYGYSVWSYCGVG 461
```

```
Qy   122 PFVNDYTWTDFSKVASGKYTANYLDFHPNELHAGDSGTFGGYPDICHDKSWDQYWLWASQ 181
         ||| |||||||||||||||||||||||||||:|      ||||| ||  :||||:||||:
Db   147 PFVGDYTWTDFSKVASGKYTANYLDFHPNEVKCCDBGTFGGFPDIAHEKEWDQHWLWASD 206

Qy   182 ESYAAYLRSIGIDAWRFPDYVKGYAPWVVKDWLNWWGGWAVGEYWDTNVDAVLNWAYSSGA 241
         |||||||||| ||||| |||||| |:|||||||||||||||||||||||||: ||||||||
Db   207 ESYAAYLRSIGVDAWRFDYVKGYGAWVVKDWLNWWGGWAVGEYWDTNVDALLNWAYSSGA 266

Qy   242 KVFDFALYYKMDEAFDNKNIPALVSALQNGQTVVSRDPFKAVTFVANHDTDIIWNKYPAY 301
         ||||| |||||||||| ||||||:||||||||||||||||||||||||||||||||||||
Db   267 KVFDFPLYYKMDEAFDNTNIPALVDALQNGGTVVSRDPFKAVTFVANHDTDIIWNKYPAY 326

Qy   302 AFILTYEGQPTIFYRDYEEWLNKDKLKNLIWIHENLAGGSTDIVYYDNDELIFVRNGYGD 361
         |||||||||| ||||||||||||| :::||||||:: |||||:|||||:::||||||||::
Db   327 AFILTYEGQPVIFYRDYEEWLNKDKLNNLIWIHDHLAGGSTSIVYYDSDELIFVRNGDSK 386

Qy   362 KPGLITYINLGSSKAGRWVTVPKFPAGACIHEYTGNLGGWVDKYVYSSGWVLEAPAYDPA 421
         :||||||||||||||:|||| |||||||||||||||||||||||||| ||||||||||||
Db   387 RPGLITYINLGSSKVGRWVYVPKFPAGACIHEYTGNLGGWVDKYVESSGWVLEAPAYDPA 446

Qy   422 NGQYGYSVWSYCGVG 436
         :|||||:||||||||
Db   447 SGQYGYTVWSYCGVG 461
```

```
RESULT 4
O33476
ID   O33476;          PRELIMINARY;       PRT;    461 AA.
AC   O33476;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Alpha-amylase precursor.
GN   APKA.
OS   Pyrococcus kodakaraensis.
OC   Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC   Thermococcus.
OX   NCBI_TaxID=69014;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KOD1;
RA   Tachibana Y., Mendez L., Fujiwara S., Takagi M., Imanaka T.;
RT   "Cloning and expression of the alpha-amylase gene from the
RT   hyperthermophilic archaeon Pyrococcus sp. KOD1, and characterization
RT   of the enzyme.";
RL   J. Ferment. Bioeng. 82:224-232(1996).
DR   EMBL; D83793; BAA21130.1; -.
DR   HSSP; P06278; 1VJS.
DR   GO; GO:0004556; F:alpha-amylase activity; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   InterPro; IPR006047; Alpha_amyl_cat.
DR   InterPro; IPR006046; Glyco_hydro_13.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   PRINTS; PR00110; ALPHAAMYLASE.
KW   Signal.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27    461       ALPHA-AMYLASE.
SQ   SEQUENCE   461 AA;  52213 MW;  FCC13LA93DC03123 CRC64;
```

```
Query Match             91.4%;  Score 2248;  DB 1;  Length 461;
Best Local Similarity   90.6%;  Pred. No. 5.9e-144;
Matches  394;  Conservative   18;  Mismatches   23;  Indels    0;  Gaps    0;

Qy     2 AKYSELEKGGVIMQAFYWDVPSGGIWWDTIRQKIPEWYDAGISAIWIPPASKGMGGAYSM 61
         ||||||| |||||||||||||| |||||||||:|||||||:|:|||||||||||||||||
Db    27 AKYSELEEGGVIMQAFYWDVPAGGIWWDTIRSKIPEWYEAGISAIWIPPASKGMGGAYSM 86

Qy    62 GYDPYDFFDLGEYDQKGTVETRFGSKQELVNMINTAHAYGMKVIADIVINHRAGGDLEWN 121
         |||||||||||||:|||||||||||||||:||||||||||:||||||||||||||||||||
Db    87 GYDPYDFFDLGEYNQKGTVETRFGSKQELINMINTAHAYGIKVIADIVINHRAGGDLEWN 146
```

Jun 30 10:27:00 2004

```
Db   1045  TACTCCGGAGCATGGCATCGACGCGCTGGGCGCTTCGACTACGTCAAGGGCTACGCTCCC 1104
           ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Qy    619  TGGGTCGTCAAGGACTGGCTGAACTCGTGGGGAGAGCTGGGGGTTGGAGAGTACTGGGAC 678
           |||||||  ||||  || ||||||||||||||||||||||||||||||||||||||||||
Db   1105  TGGGTCGTTAAGAACTCGCTGAACCCGGTCGGGGCGGTTGGGGGTTGGAGAGTACTGGGAC 1164
Qy    679  ACCAAGTCGCACGCTGTTCTCAACTGGGCATACTCGAGCGGTGCCAAGGTCTTTGACTTC 738
           ||||||||||| || ||||||||||||||||||||||||||||||||||| ||||||||||
Db   1165  ACCAAGTCGATGGCACTCCCTGAGCTGGGCCTACGACAGCGGTGCTAAAGTCTTCGACTTC 1224
Qy    739  GCCCTCTACTACAAGATGATGAGGCCTTTGACAACAAAACATTCCAGCCGCTCGTCTCT 798
           |||||||||||||| || ||||| ||||  ||||||||  ||| || || ||||||||
Db   1225  CCGCTCTACTACAAGATGACGAGGGCCTTCGATAAGTAACAACATCCCGCCCGTCGTGGAC 1284
Qy    799  GCCCTTCAGAACGGCCAGACTGTTGTCTCCCCGGACCCGGTTCAAGGGCGTAACCTTTGTA 858
           ||||||||||||  ||| |||||| |||  |||||||| ||||| || || || |||||
Db   1285  GCCCTCAAGAACGGAGGCACGGGTCGTACGCCCGGACCCCGTTCAAAGCCGTGACCTTCGTT 1344
Qy    859  GCAAACCACGACCACCGATATAATCTGGAACAAGTATCCAGCTCCAGCTCCTCATCCTCACC 918
           ||  ||||| ||||| ||||| ||| ||||||||||||  |||||||||||||||||||||
Db   1345  GCCAACCACGATCCAACATAATCTGGAACAAGTATCCGGCCTACGCCTACGCCTTCATCCTCACC 1404
Qy    919  TACGAGGGCCAGCCGACAATATTCTACCGGACTACAGGAGGTGGCTCAACAAGGATAAG 978
           || || || ||| || ||||| ||||||||||| |||||  ||||||||||||| |||
Db   1405  TATGAGGGACAGCCGCAATATTCTACCGGACTACAGGAGGTGGCTCAACAAGGACAGG 1464
Qy    979  CTCAAGAACCTCATCTGGATACATGAGAACCTCGCCGAGGAGGAAGCACCGACATAGTCTAC 1038
           || ||||||||||||| |||  |||||||  ||||||||||||||||||||| ||| ||||
Db   1465  CTCAGGAACCTCATCTCTGATACACGACCACCTCGCGGGAGGAAGCACAGACATCATCTAC 1524
Qy   1039  TACGATAACGATGACTCATCTTCGTCAGGAACGGCTCAGGGGACAAGCCGCGGGGCTTATA 1098
           |||| |||||||| |||||||||||||||||||  || ||||| ||| || |||||
Db   1525  TACGACAGGGACGAGCTTATCTTCGTCGGAGAAAACGGCTACGGGAGACAAGCCGGGACTGATA 1584
Qy   1099  ACCTACCATCAACCTAGGCTCGAGCAAGCGCCGAAGGTGGGTTTATGTCCGGAAGTTCGCG 1158
           |||||||||||| ||| |||  ||||||| |||||| |  |||| ||  ||||||| ||||
Db   1585  ACCTACCATCAACCTCGGGCTCAAGCAAGCCGGAGGTGGGTCTACGTTCCGAAGTTCGCA 1644
Qy   1159  GGCGCGGTGCATCCACGAGTATACTGGTAACCTCGGAGGCTGGGTAGACAAGTAGTCTAC 1218
           ||||| ||| |||| ||||| |||| || ||| ||| ||||| || ||||| ||||
Db   1645  GGCTCGTGCATACACGAGTACACCGGGCAACCTCGGCGGCGGATTGACAAGTGGGTTGAC 1704
Qy   1219  TCAAGCGGCTGGGTCTATCTGAAGCTCCAGCTTACGACCCTGCCAACCGGCCAGTATGGC 1278
           |||| || ||||| |||| || ||||||||||| ||||| |||||||||| || |||||
Db   1705  TCAAGCGGTCGGGTCTACCTTGAGCCCTACCGGCCCAACGCGACCCCGACCGGCCAGTACGGC 1764
Qy   1279  TACTCCGTGTGGAGCTACTGCGGGGGTGGGCTGA 1311
           |||||| ||||||||||||| ||||||| |||||
Db   1765  TACTCCGTTTGGAGCTACTGCGGGGGTGGGCTGA 1797
```

FEATU   1 and Molecular
                                         ims-Champagne-Ardenne,
                                         51687, France

                                         lis"

PAKAETLENGGVIMQAFYWDVPG
SMGYDPYDFFDLGEYYQKGSVE
DLEWNPFTNSYTWTDFSKVASGK
IWLWASNESYAAYLRSIGIDAWR
SWAYDSGAKVFDPLYYKMDEA
NIIWNKYPAYAFILTYEGQPAI
DELIFVRNGYGDKPGLITYINL
uSGRVYLEAPAHDPANGQYGYSV

ORIGIN   WSYCGVG"

Query Match   76.5%;   Score 1003.4;   DB 1;   Length 2705;
Best Local Similarity   86.0%;   Pred. No. 3.4e-216;
Matches 1112;   Conservative   0;   Mismatches 181;   Indels 0;   Gaps   0;

```
Qy     19  CTGGAAAAGGGCGGGTCATAATGCAGGCGTTCTCTACTGGGACTGCCTTCAGGAGGAATA 78
           |||||| || ||||| ||||||||||||| |||||||||||||||||||||||| |||
Db    505  CTTGAGAACGGCGGCGTCATAATGCAGGCGTTCTACTGGGACGTCCCAGGTGGAGGAATC 564
Qy     79  TGGTGGGACACAATACGGCGAGAGATACCGGAGTGGTACGATGCCGGAATCTCCGCAATA 138
           ||||||||||  || ||| |||||||  ||  |||||| || ||||| ||  ||||||||
Db    565  TGGTGGGACACCATAGCCCAGAAGATACCCCGACTGGGGCGCCGGAGCGCCGGGGATTTCGGCAATA 624
Qy    139  TGGATTCCCCGGCGGAGCAAGGGCATGGGCGCCGCCTATTCGATGGGCTACGACCCCTAC 198
           ||||||||||| |||||||||| |||||| ||||| ||||||||||| ||||||||| ||
Db    625  TGGATTCCTCCCGCCAGTAGGCATGAGCGCGCGCGGCTATTCGATGGGCTACGACCCCCTAC 684
Qy    199  GACTTCTTTGACCTCGGTGAGTACGACCAGAGAAGGAAGCGTAGAAGCGCGCTTTGGCTCC 258
           |||||||| ||||||| || ||| |||| |||| |||||||||  |||| |||||||||||
Db    685  GATTTCTTCGAACCTCGGTGAGTACTACCAGAGAAGGGAAGCGTTGAAGACCCGCTTCGGATCA 744
Qy    259  AAGCAGGAGCTCGTGAACATGATAAACACCGCCCACGCCCATGCGCATGAAGGTAATAGCC 318
           ||||||||||| |||||||||||||  ||| ||||||||||||| || |||||||||| ||
Db    745  AAAGAGGAGCTTGTGAACATGATAAACACCGCCCATGCTCACAACATGAAGGTCATAGCG 804
Qy    319  GATATAGTCATCAACCACCGCGCCGGCGGTGGACCTTGGAGTGGAACCCCTTCGTGAACGAC 378
           |||| ||||||||||  ||| ||||| || |||||||||| || ||| || ||||| ||||
Db    805  GACATAGTCATCAACCACCGCGCCGGCGGATTCGAGTGGAATCCTTCACCAACAGC 864
Qy    379  TATACCTGGACCGACTTCTCAAAGGTCGCGTCGGTAAATACACGGCCAACTACCTCGAC 438
           || |||||||| ||| || |||||||  ||| |||| || ||||||||||||||||||
Db    865  TACACCTGGACCGATTTCTCTGAAGGTCGCGTGGGGCAAGCAACTACACGGCCAACTACCTCGAC 924
Qy    439  TTCACCCGAACGAGCTCCATGCGGGCGGATTCCGGAACATTTCGAGGGCTATCCGACATA 498
           ||||| || ||||||| ||  ||||||| |||||||||| ||||||||||||||||||||
Db    925  TTCCACCCGGAACGAGCTTCACGCGGCGATTCCGGAACATTTGGAGGGCTATCCCGACATA 984
Qy    499  TGCCACGACAAGAGAGCTGGGACCAGTACTGGCGCTTCGGACTTCGGCCTTCGACACTGCTCCC 558
           ||||| ||| |||||||||||| || || |||||| ||||||||||||||| |||||||
Db    985  TGCCACGACAAGAGAGCTGGGACCAGTACTGGCCGCAGCGAAAGCTACGCCGCC 1044
Qy    559  TAACTCAGGAGCATCGAGTGCGATCGATGCCTAGCCGCTTCGGACTACGTCAAGGGCTATGCTCCC 618
```

us-10-081-739a-1.rge

10:27:00 2004

Faculty of Engineering, Biotechnology; 2-1, Yamadaoka, Suita, Osaka
565, Japan (Tel:06-879-7442, Fax:06-879-7448)
```
FEATURES          Location/Qualifiers
   source         1..2179
                  /organism="Pyrococcus sp."
                  /mol_type="genomic DNA"
                  /strain="KOD1"
                  /db_xref="taxon:33866"
   gene           452..1837
                  /gene="ApkA"
   CDS            452..1837
                  /gene="ApkA"
                  /codon_start=1
                  /transl_table=11
                  /product="alpha-amylase"
                  /protein_id="BAA21130.1"
                  /db_xref="GI:2251108"
                  /translation="MKKFVALLLIMFFVVSMAAVAQPASAAKYSELEEGGVIMQAFYW
                  DVPAGGIWWDTIRSKIPEWYEAGISAIWIPPASKGMGGAYSMGYDPYDFFDLGEYNQK
                  GTVETRFGSKQELINMINTAHAYGIKVIADIVINHRAGGDLEWNPFVGDYTWTDFSKV
                  ASGKYTANYLDFHPNEVKCCDEGTFGGPDIAHEKEWDQHWLWASDESYAAYLRSIGV
                  DAWRFDYVKGYGAWVVKDWLNWWGGWAVGEYWDTNVDALLNWAYSSGAKVFDFPLYYK
                  MDEAFDNTNIPALVDALQNGGTVVSRDPFKAVTFVANHDTDIIWNKYPAYAFILTYEG
                  QPVIFYRDYEWLNKDKLNNLIWIHDHLAGGSTSIVYYDSDELIFVRNGDSKRPGLIT
                  YINLGSSKVGRWVVPKFAGACIHEYTGNLGGWHVDKYVESSSWVYLEAPAYDPASGQY
                  GYTVWSYCVG"
   sig_peptide    452..529
                  /gene="ApkA"
   mat_peptide    530..1834
                  /gene="ApkA"
                  /product="alpha-amylase"
ORIGIN

Query Match        75.6%;  Score 991.2;  DB 1;  Length 2179;
Best Local Similarity 84.9%;  Pred. No. 2e-213;
Matches 1110;  Conservative   0;  Mismatches 198;  Indels   0;  Gaps   0;

Qy        4 GCCAAGTACTCCGAGCTGGAAAAGGGCGGGGTCATAATGCAGAGGCGTTCATACTGGGACGTG 63
            |||||| || ||||||| |||| |||||||| |||||  || ||||| | |||||||||||||
Db      530 GCAAAGTATTCCGAACTCGAAGAGAGGGGCGGCGTTATAATGCAGGCCTTCACTACTGGGACGTT 589

Qy       64 CCTTCAGGAGGAATAAGTTGTGGGACACAAATACGGCAGAGATAACCGGAGTCGTAGCGATGCC 123
            ||| |||||||| || ||| || |||||||||| | ||||| || || |||||||||||||
Db      590 CCAGCGGGGAGGAATCGTGTGGGGATACAATCAGAGCCAAGCAAGATACCCGAGTGTAGGAGGCT 649

Qy      124 GGAATCTCCGCAATATGAATTCCCCGGCGAGCCAAGGGCATGGGCCGGCCGCGCTATTCGATG 183
            |||||||||| ||||| ||||||||||||||||||||  ||||| ||||||||||||||||||
Db      650 GGAATCTCCGCCATCGGATTCCCGCCAGCCAGCAAGGGCATGGGAGGAGCTTATTCAATG 709

Qy      184 GGCTACGACCCCTACGACTTCTTTGACCTTGACCTCGGTGGTGAGTCGGTGGAACGGTAGAG 243
            |||||||| |||||||||||||| || |||||||||||||||||| |||| ||| |||||||
Db      710 GGCTACGACCCCATACGACTTCTTCGACCTCGGACCTCGGACCTCTCCGGAGTCAACCAGAACAGTTGAA 769

Qy      244 ACGGCTTGGCTCCAAGCAGAGGAGCTCGTGAACATGATAAACACCGCCCAGCCTATGGC 303
            || |||||||| |||||| |||||||||||||| |||||| ||||||||||||||||||
Db      770 ACTCGCTTTGGCTCAAAGCAGAGGAGCTTATCAACATGATAAACACCGCCCATGCCTACGGC 829

Qy      304 ATGAAGGTAATAGCCGATATAGTCATCAACAACCACCGCGCCGGTGACTTGGAGTCGGAAC 363
            |||||| || ||||||||||| |||||||||||||||| || |||||| |||| ||||||
Db      830 ATAAAGGTCATAGCTGATATCGTTATCAACATCCATCGGGGCCGAGGGATCTCGGGTAAATACG 889

Qy      364 GCCAACTACCTCGACTTCCACCCGAACGAGCTTCAAAGTCGCGTCGGGTAAATACACG 423
            |||||||||||||||||||| ||| |||| |||||||||| || ||||||||||| |
Db      890 GCCAACTACCTCGACTTCCACCCCAACGAGGTGAAGGTGCCTCGGGCAAATATACC 949

Qy      424 GGCTATCCCGACATATGCCACGACGACAAGATGCCACAGTACTGGGACCTGCTTCGGCCAG 483
            || || ||||| || ||| ||||||||||||| |||||||||||||||| ||||||||||
Db      950 GCCAACTACCTCGACTTCCACCCCCACGAGCTCAAGTGCTGTGACGAGGGCACCATTTGGA 1009

Qy      484 GGCTATCCCGACATATGCCACGACGACAAGATGCCACAGTACTGGGACCTGCTTCGGCCAG 543
            || ||| |||| || ||| || ||||||||| |||||||||||||||| ||||||| || |
Db     1010 GGTTTCCCAGACATAGCCCACGAGAAGGAGTGGGACCAGCACTGGCTCTGGGCGAGCGAC 1069
```